REMARKS

In view of the following discussion, the Applicant submits that none of the claims now pending in the application is made obvious under the provisions of 35 U.S.C. §103. Thus, the Applicant believes that all of these claims are now in allowable form.

I. REJECTION OF CLAIMS 1-6 AND 8-36 UNDER 35 U.S.C. §103

The Examiner rejected claims 1-6 and 8-36 under 35 U.S.C. §103(a) as being unpatentable over the Brown et al. patent (U.S. Patent No. 5,719,997, issued February 17, 1998, hereinafter "Brown") in view the Ehsani et al. application (U.S. Publication No. 2002/0032564, published March 14, 2002, hereinafter "Ehsani"). In response, the Applicant has amended independent claims 1, 11, 18, 34 and 35 from which claims 2-6, 8-10, 12-17 and 19-33 depend, in order to more clearly recite aspects of the invention. The rejection is respectfully traversed.

Particularly, the Examiner's attention is directed to the fact that Brown and Ehsani both fail to disclose or suggest the novel invention of generating a grammar and one or more related subgrammars (including, for example, a word subgrammar, a phone subgrammar and a state subgrammar) based at least in part on a grammar provided by a remote computer, as claimed in Applicant's independent claims 1, 11, 18, 34, 35 and 36.

In contrast, Brown teaches a method in which a speech recognition system possesses an entire grammar, but merely instantiates selected portions of the grammar over time, *e.g.*, as more of the incoming speech signal is received. Specifically, as the Examiner concedes on page 5 of the Office Action, "Brown does not explicitly teach that the set of data structures [*i.e.*, the top-level grammar and associated subgrammars] is sent through a communication channel by a remote computer, or selected thereby or that the set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer" (emphasis added). In other words, Brown does not teach, show or suggest the desirability of distributing the top-level grammar and related subgrammars, or the information necessary to generate the grammars, in order to conserve memory at the speech recognition system.

Ehsani does not bridge this gap in the teachings of Brown. Specifically, Ehsani also does not teach, show or suggest acquiring a set of data structures <u>including a grammar, a word subgrammar, a phone subgrammar and a state subgrammar</u>, where the data structures are generated at least in part based on <u>a grammar provided by a remote computer</u>. The portion of Ehsani that the Examiner cites to support this limitation does not, in fact, teach that a speech recognition system acquires a grammar or set of grammars from a remote computer. Rather, the cited portion of Ehsani teaches that a user may access an application or database remotely via a voice telephony server (see, Ehsani, paragraph [0200]: "[C]allers dial into a voice telephony server and are led through a series of voice-driven interactions that lets them complete automated transactions such as getting information, accessing a database or making a purchase"). In other words, Ehsani teaches that the <u>speech signal or input</u> to be processed (*i.e.*, the caller's voice commands) may be provided remotely to the speech recognition system (*e.g.*, via a telephone or handheld device). In addition, the application controlled by the voice telephony server may reside on a separate server. However, the <u>grammar</u> used to process the remotely provided speech signals resides <u>locally, on the voice telephony server/speech recognition system</u>. <u>Nowhere</u> does Ehsani teach that <u>a grammar</u> for generating a set of data structures used by the telephony server to process the voice commands is <u>provided by a remote computer or server</u>.

Brown in view of Ehsani thus fails to disclose or suggest the novel invention of <u>generating a grammar and one or more related subgrammars</u> based at least in part on a grammar provided <u>by a remote computer</u>, as claimed in Applicant's independent claims 1, 11, 18, 34, 35 and 36. Specifically, Applicant's claims 1, 11, 18, 34, 35 and 36 positively recite:

1. A method for allocating memory in a speech recognition system comprising the steps of:

acquiring a first set of data structures <u>that contain a grammar, a word subgrammar, a phone subgrammar and a state subgrammar</u>, each of the subgrammars related to the grammar, <u>wherein the first set of data structures is generated by the speech recognition system based at least in part on a grammar provided by a remote</u>

computer;

acquiring a speech signal;

performing a probabilistic search using the speech signal as an input, and using the first set of data structures as possible inputs; and

allocating memory for one of the subgrammars when a transition to that subgrammar is made during the probabilistic search. (Emphasis added)

11. In a speech recognition system, a method for recognizing speech comprising the steps of:

acquiring a first set of data structures that contain a grammar, a word subgrammar, a phone subgrammar and a state subgrammar, each of the subgrammars related to the grammar, wherein the first set of data structures is generated by the speech recognition system based at least in part on a grammar provided by a remote computer;

acquiring a speech signal;

performing a probabilistic search using the speech signal as an input, and using the first set of data structures as possible inputs;

allocating memory for one of the subgrammars when a transition to that subgrammar is made during the probabilistic search; and

computing a probability of a match between the speech signal and an element of the subgrammar for which memory has been allocated. (Emphasis added)

18. In a speech recognition system, a method for recognizing speech comprising the steps of:

acquiring a first set of data structures that contain a top level grammar and a plurality subgrammars, each of the subgrammars hierarchically related to the grammar and to each other, wherein the first set of data structures is generated by the speech recognition system based at least in part on a grammar provided by a remote computer;

acquiring a speech signal;

performing a probabilistic search using the speech signal as an input, and using the first set of data structures as possible inputs;

allocating memory for specific subgrammars when transitions to those specific subgrammars are made during the probabilistic search; and

computing probabilities of matches between the speech signal and elements of the subgrammars for which memory has been allocated. (Emphasis added)

34. A method for allocating memory in a speech recognition system comprising the steps of:

acquiring a set of data structures that contain a grammar and one or more subgrammars related to the grammar, wherein the first set of data structures is generated by the speech recognition system based at least in part on a grammar

provided by a remote computer;

acquiring a speech signal;

performing a probabilistic search using the speech signal as an input, and using the set of data structures as possible inputs; and

allocating memory for a selected one or more of the subgrammars when a transition to the selected subgrammar is made during the probabilistic search. (Emphasis added)

35. In a speech recognition system, a method for recognizing speech comprising the steps of:

(a) acquiring a set of data structures that contain a grammar and one or more subgrammars related to the grammar, wherein the first set of data structures is generated by the speech recognition system based at least in part on a grammar provided by a remote computer;

(b) receiving spoken input;

(c) using one or more of the data structures to recognize the spoken input;

(d) while the speech recognition system is operating, acquiring a second set of data structures that contain a second grammar and one or more subgrammars related to the second grammar; and

(e) repeating steps (b) and (c), using the second set of data structures in step (c). (Emphasis added)

36. In a speech recognition system, a method for recognizing speech comprising the steps of:

(a) acquiring from a first remote computer a set of data structures that contain a grammar and one or more subgrammars related to the grammar;

(b) receiving spoken input;

(c) using one or more of the data structures to recognize the spoken input;

(d) while the speech recognition system is operating, acquiring a second set of data structures from the first remote computer or from a second remote computer, the second set of data structures containing a second grammar and one or more subgrammars related to the second grammar; and

(e) repeating steps (b) and (c), using the second set of data structures in step (c). (Emphasis added)

Applicant's invention is directed to a method for allocating memory in a speech recognition system. Conventional speech recognition systems require a great deal of memory in order to accommodate and process large vocabularies. These systems typically compile, expand, flatten and optimize all grammars contained in a system

vocabulary into a large, single-level data structure that must be stored in memory before the speech recognition system can operate. Such techniques substantially restrict the capabilities of speech recognition systems that operate on limited memory and processing power, such as portable speech recognition systems.

The present invention provides a method for speech recognition in which memory is allocated to a particular system subgrammar when a transition is made to that subgrammar during a probabilistic search. A system vocabulary has a hierarchical data structure including at least one top-level grammar (*e.g.*, "Days of the Week") and at least one subgrammar within that top-level grammar such as a word subgrammar (*e.g.*, Monday, Tuesday, Wednesday, etc.), a phone subgrammar (*e.g.*, /m/, /ah/, /n/, /d/, /ey/, etc.) and a state subgrammar (*e.g.*, comprising Hidden Markov Models). When the system receives a speech signal for processing, the speech signal is input, along with the (unexpanded) top-level grammar and one or more subgrammars, into a probabilistic search. When a transition is made to a particular subgrammar during the probabilistic search, memory is allocated to the subgrammar, which may then be expanded and evaluated to assess the probability of a match between the speech signal and an element in the subgrammar. In this manner, memory is conserved and allocated only to portions of the system vocabulary that are currently needed for speech processing. In addition, at least part of the information used to generate the top-level grammar and/or the related subgrammars (*e.g.*, a selected grammar) may be provided (or selected from a set of local possibilities) by a remote computer or server, to further conserve the memory required to operate the speech recognition system (which may be implemented, for example, in a portable device).

In contrast, Brown only teaches a speech recognition system that uses an evolutional grammar to recognize an input speech signal in real time. In particular, Brown teaches that as speech recognition processing begins, only a portion of a system grammar (*i.e.*, a vocabulary comprising a plurality of interrelated words) is implemented for recognition purposes. As more of the speech signal is received by the system and as processing proceeds, additional portions of the grammar network (*i.e.*, additional words or vocabulary) are implemented as necessary. In other words, a single system

grammar is assembled, piece-by-piece, as the speech signal is received. As discussed above, Brown clearly fails to disclose or suggest the novel invention of generating a grammar and one or more related subgrammars based at least in part on a grammar provided by a remote computer (*e.g.*, a grammar provided directly by the remote computer, or a local grammar that is selected by the remote computer).

Ehsani teaches a method for creating grammar networks for use in natural language voice user interfaces (NLVUIs). Valid phrases are extracted from a text corpus and clustered into classes to create a "thesaurus" of fixed word combinations that represent different ways of saying the same thing. In this way, anticipated user responses can be expanded into alternative linguistic variants. Like Brown, Ehsani also fails to disclose or suggest the novel invention of generating a grammar and one or more related subgrammars based at least in part on a grammar provided by a remote computer.

Thus, Brown in view of Ehsani fails to disclose or suggest the novel invention of generating a grammar and one or more related subgrammars based at least in part on a grammar provided by a remote computer, as claimed in Applicant's independent claims 1, 11, 18, 34, 35 and 36. Therefore, the Applicant submits that independent claims 1, 11, 18, 34, 35 and 36 fully satisfy the requirements of 35 U.S.C. §103 and are patentable thereunder.

Dependent claims 2-6, 8-10, 12-17 and 19-33 depend from claims 1, 11 and 18 and recite additional features therefore. As such, and for at least the same reasons set forth above, the Applicant submits that claims 2-6, 8-10, 12-17 and 19-33 are not made obvious by the teachings of Brown in view of Ehsani. Therefore, the Applicant submits that dependent claims 2-6, 8-10, 12-17 and 19-33 also fully satisfy the requirements of 35 U.S.C. §103 and are patentable thereunder.

II. CONCLUSION

Thus, the Applicant submits that all of the presented claims now fully satisfy the requirements of 35 U.S.C. §103. Consequently, the Applicant believes that all of these claims are presently in condition for allowance. Accordingly, both reconsideration of this

application and its swift passage to issue are earnestly solicited.

If, however, the Examiner believes that there are any unresolved issues requiring the issuance of a final action in any of the claims now pending in the application, it is requested that the Examiner telephone Mr. Kin-Wah Tong, Esq. at (732) 530-9404 so that appropriate arrangements can be made for resolving such issues as expeditiously as possible.

Respectfully submitted,

1/5/06
Date

Kin-Wah Tong, Attorney
Reg. No. 39,400
(732) 530-9404

Patterson & Sheridan, LLP
595 Shrewsbury Avenue
Shrewsbury, New Jersey 07702